



NICOLAS

ETOP 6-14

**A phase II trial evaluating the safety and efficacy of the addition of
concurrent anti-PD-1 nivolumab to
standard first-line chemotherapy and radiotherapy in locally advanced
stage IIIA/B Non-Small Cell Lung Carcinoma**

Statistical Analysis Plan (SAP) Final efficacy analysis

A clinical trial of ETOP

SAP VERSION	V.1.1
SAP VERSION DATE	10 June 2019

Contents

INTRODUCTION	4
1 Trial oversight (as per protocol - Version 3.0).....	5
1.1 Objectives.....	5
1.2 Endpoints	5
1.3 Eligibility criteria.....	6
1.4 Previous protocol versions.....	6
1.5 Statistical design, sample size & power	7
1.6 Total trial duration	8
2 General considerations	9
2.1 Analysis Timing.....	9
2.2 Data Retrieval Information.....	9
3 Statistical considerations	10
3.1 Study's endpoints.....	10
3.1.1 Primary safety endpoint (previously analysed).....	10
3.1.2 Key-secondary efficacy endpoint	10
3.1.3 Other secondary endpoints.....	10
3.2 (Serious) Adverse Events.....	11
3.3 Analysis populations	12
3.3.1 Interim safety analysis cohort (applicable for previous formal safety analysis).....	12
3.3.2 Safety cohort.....	12
3.3.3 Efficacy cohort.....	12
4 Study analysis.....	13
4.1 Patient accrual & Baseline characteristics	13
4.2 Treatment administration and follow-up information	13
4.3 Efficacy analysis.....	14
4.3.1 Key-secondary endpoint: One-year PFS.....	14
4.3.2 Further efficacy analysis based on PFS	14
4.3.3 Analyses of other secondary efficacy endpoints	15

4.3.4	Subgroup analysis.....	15
4.3.5	Exploratory analysis for patients receiving nivolumab	15
4.4	Safety analysis.....	16
4.5	Missing Data	17
4.6	Sensitivity analysis.....	17
4.7	Presentation of results.....	18
5	Technical details.....	19

INTRODUCTION

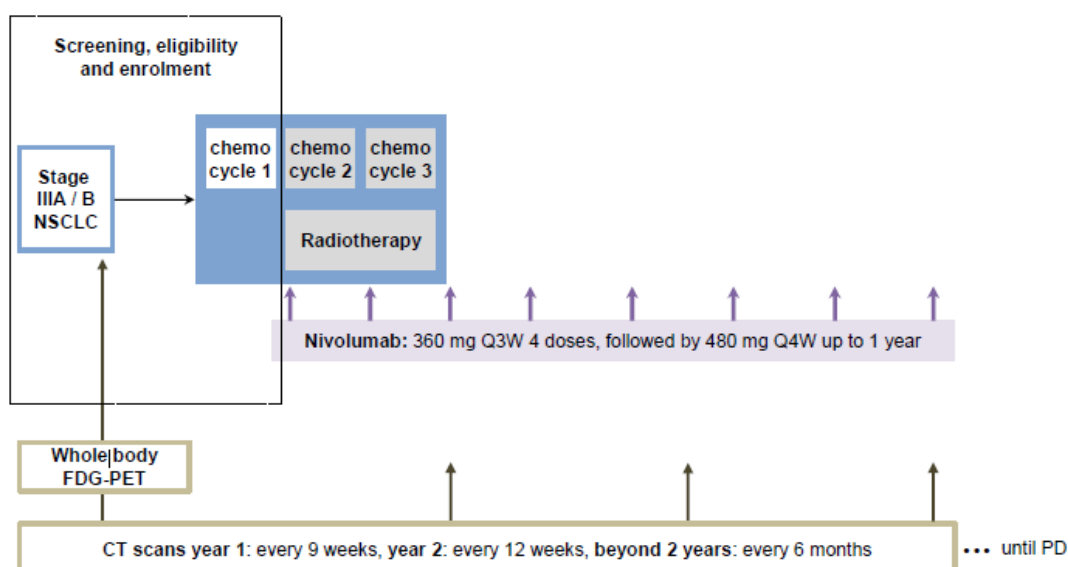
The aim of this Statistical Analysis Plan (SAP) is to describe an analytic and solid framework that will be followed in order for the final efficacy analysis of the NICOLAS trial to be implemented (based on NICOLAS protocol version 3; protocol amendment-2 [AM2]).

A short description of the contents of this statistical analysis plan is provided below:

1. **NICOLAS trial outline:** trial's schema, eligibility criteria, objectives, trial duration, previous protocol versions, sample size & power
2. **General considerations:** Analysis timing, data retrieval
3. **Statistical considerations:** definition of primary and secondary endpoints, (serious) adverse events definition, analysis populations
4. **Study Analysis:** Baseline characteristics, treatment administration and follow-up, efficacy analysis, safety analysis, handling of missing data
5. **Technical** issues, including testing and reporting conventions.
6. **List of Tables and Figures**

1 Trial oversight (as per protocol - Version 3.0)

NICOLAS is a single-arm multicenter phase II trial, evaluating the safety and efficacy of the addition of concurrent anti-PD-1 nivolumab to standard first-line chemotherapy and radiotherapy in locally advanced stage IIIA/B Non-Small Cell Lung Carcinoma.



SCHEMA 1. Trial design under protocol V3.0

1.1 Objectives

Primary objective

The **primary objective** of the study is to assess the safety and efficacy of the concurrent nivolumab administration with standard first-line chemotherapy and radiotherapy in locally advanced stage IIIA/B NSCLC, as defined by the rate of grade ≥ 3 pneumonitis (CTCAE V4.0) 6 months post-radiotherapy and, if safety is proven, to assess the progression-free survival.

Secondary objectives

- The **secondary objectives** of the study include the following measures of clinical efficacy: time to first grade ≥ 3 pneumonitis (TFP3), objective response rate (ORR), time to treatment failure (TTF) and overall survival (OS).
- To assess the safety and the tolerability of the treatment.

1.2 Endpoints

Primary endpoint:

- Grade ≥ 3 **pneumonitis** (CTCAE V4.0) observed any time during 6 months from end of radiotherapy.

Key secondary endpoint (if valid):

- One-year progression-free survival by RECIST v1.1

Secondary endpoints:

- Time to first grade \geq 3 pneumonitis
- Objective response determined by RECIST v1.1
- Time to treatment failure
- Overall survival
- Adverse events graded according to CTCAE V4.0

1.3 Eligibility criteria

Inclusion criteria at enrolment (briefly stated; for more details, refer to trial protocol):

- Histologically or cytologically confirmed locally advanced stage IIIA or III B (T0-3 N2-3 or T4 N0-3 M0) non-small cell lung carcinoma (NSCLC), according to 7th TNM classification
- Nodal status N2 or N3 need to be proven (by biopsy, EBUS, mediastinoscopy or thoracoscopy) except for overt cT4 disease
- Measurable disease according to RECIST v1.1
- Previous delivery of a maximum of one 3-weekly cycle of platinum-based chemotherapy
- ECOG performance status 0-1
- Adequate hepatic, hematological and renal function
- All AEs from previous therapies (including the first chemotherapy cycle in the context of this trial) resolved to grade <2 (except fatigue, alopecia, nausea lack of appetite or peripheral neuropathy)

Exclusion criteria at enrolment (briefly stated; for more details, refer to trial protocol):

- Metastatic disease (as determined by PET-CT and brain MRI (preferred) or high-quality brain CT with intravenous contrast at the time of staging, performed within 35 days before the beginning of first chemotherapy cycle)
- Previous radiotherapy to the chest, including radiotherapy for breast cancer
- Prior chemotherapy, radiotherapy or molecular targeted therapy for NSCLC (with the exception of one cycle of chemotherapy given prior to enrolment into this trial)
- Active, known or suspected autoimmune disease

1.4 Previous protocol versions

In the original protocol, radiotherapy (RT) (≥ 60 Gy), was delivered either concurrently with chemotherapy in cycles 2 and 3 or sequentially, while nivolumab administration started after completion of the chemo-radiotherapy phase. Two protocol amendments have been approved (AM1, AM2).

In AM1, enrollment was allowed by cycle 2, while Nivolumab administration was starting concurrently with RT.

In AM2, only concurrent chemo-radiotherapy was allowed (sequential chemo-RT was not allowed anymore). Chemotherapy was administered in 3 cycles, and nivolumab first in four doses (360 mg) every 3 weeks and thereafter, every 4 weeks (480 mg) for up to one year.

The main change that occurred in AM2, has been the adaptation of the statistical design (and corresponding increase of sample size), so as to facilitate, beyond the safety evaluation, the efficacy evaluation of trial treatment. The key-secondary efficacy endpoint (one-year PFS rate), has been introduced in AM2, and is designed to be hierarchically tested, conditional on proven adequate safety, as described in more detail in the next section.

1.5 Statistical design, sample size & power

A **total sample size of 78 patients**, receiving concurrent therapy is needed for the study's objectives, assuming a 5% loss-of-follow-up rate. First, testing for safety is designed to evaluate the 6-month post-radiotherapy pneumonitis-free rate and if the safety endpoint is met, an efficacy analysis will be performed regarding the one-year progression-free survival (PFS) (hierarchical design).

More particularly, the **primary safety hypothesis** corresponds to the 6-month pneumonitis-free rate of grade ≥ 3 : Null hypothesis states that 6-month pneumonitis-free rate of grade ≥ 3 is less than or equal to $\pi_0=67\%$, vs the one-sided alternative that the rate is above 67%, tested at $\pi_1=85\%$:

$$H_0: \pi_0 \leq 67\% \text{ vs } H_1: \pi_1 > \pi_0, \text{ at } \pi_1 = 85\%$$

For a one-sided alpha 0.05 and power 83%, the required sample size for the safety evaluation is 41 evaluable patients allowing for one interim safety analysis at 21 patients without any requirement for trial interruption. In the safety evaluation phase of the trial 43 patients will be recruited allowing for around 5% competing risk rate. If additional patients are non-evaluable for the primary endpoint, up to 3 patients will be replaced.

An **interim safety analysis** is designed to be performed when 21 patients have completed a 3-month follow-up on nivolumab after chemotherapy and radiotherapy, assuming approximately 70% of the cases of pneumonitis occur by 3 months after the end of radiotherapy. If at the safety interim analysis, none of the 21 patients has developed pneumonitis of grade ≥ 3 by 3 months, then the safety phase of the trial could stop early with the conclusion that the treatment is feasible and safe. The trial will continue to reach the total sample size required to test the efficacy hypothesis. If the only patients developing pneumonitis of grade ≥ 3 by 3 months have done so before starting nivolumab, the Independent Data Monitoring Committee (IDMC) could choose to focus on the patients who started nivolumab for claiming feasibility.

If the trial does not reach the safety conclusion at the interim, the final safety analysis will be performed when 41 patients have completed the 6-month follow-up. If at least 33 out of 41 patients reach 6 months pneumonitis-free, the 6-month pneumonitis-free rate will be considered as promising, and the trial will continue to allow for the treatment to be tested for efficacy.

Based on a hierarchical design, a **key-secondary efficacy hypothesis** (secondary- H_s) is tested after the pneumonitis null hypothesis (primary- H_p) is rejected. The one-year PFS rate that could be achieved with current best provided care is estimated to be around 45%. The aim of the combination under investigation is to improve the one-year PFS by at least 15%, that is, to achieve a one-year PFS rate of at least 60%.

Efficacy hypothesis (H_s): H_{s0} : $PF_{s0} \leq 45\%$ vs H_{s1} : $PF_{s1} > PF_{s0}$, at $PF_{s1}=60\%$

A sample size of 74 evaluable patients provides 83% power, for testing this efficacy hypothesis, based on the exact binomial test for single proportion, at the one-sided alpha of 0.05. Assuming 5% non-evaluable patients, a total sample size of 78 patients need to enter the study.

The final efficacy hypothesis is tested after the safety null hypothesis is rejected either at the interim or the final safety analysis, when 74 evaluable patients under concurrent chemo-radiotherapy have reached one-year follow-up (from time of enrolment into the trial) or have experienced a PFS event.

The 78 patients will be enrolled under the protocol amendments (protocol versions 2.0 and 3.0). Patients included under the original protocol (version 1.0) will be evaluated separately.

1.6 Total trial duration

After a run-in period of 3 months for the activation of the centres, patient accrual is expected to be completed within 9 months (under AM2 and taking into account the accrual accomplished under AM1).

The study duration for the key secondary efficacy endpoint will be approximately 22 months, including a 3-month start-up period, 7 months accrual (of the additional patients) and 12 months follow-up for evaluating the one-year PFS rate.

Follow-up will continue until 2 years from start of nivolumab treatment of the last recruited patient. The trial will end with the preparation of the final report, scheduled at 2.5 years after the inclusion of the first patient.

2 General considerations

2.1 Analysis Timing

According to the statistical design, the final efficacy analysis (key secondary endpoint) is scheduled to be performed when the last enrolled patient has completed 1 year on follow-up (and given that the safety hypothesis has been already proven).

2.2 Data Retrieval Information

The final analysis will be based on the final database download, to be performed as soon as all patients have completed one-year of follow-up or have a PFS event before that.

Based on this database extraction, a series of queries will be created in order to clean the data as much as possible. The queries will be forwarded to the data manager of the protocol and a prespecified period of time will be given for the queries to be answered. Corrections and responses based on these queries, will be used for correcting the previously downloaded database.

3 Statistical considerations

3.1 Study's endpoints

3.1.1 Primary safety endpoint (previously analysed)

The primary (safety) endpoint is grade ≥ 3 pneumonitis (CTCAE v4.0) observed any time during 6 months from the end of radiotherapy. It is defined as the number of patients reaching up to 6 months post chemo-radiation treatment without any episode of CTCAE v4.0 grade ≥ 3 pneumonitis.

3.1.2 Key-secondary efficacy endpoint

Progression-free survival (PFS) is the key secondary (efficacy) endpoint. It is defined as the time from the date of enrolment until documented progression or death, if progression is not documented. Censoring will occur at the last tumour assessment.

Of note, patients who continue treatment beyond initial investigator-assessed, RECIST v1.1-defined progression will be considered to have had progressive disease at the time of the initial progression event.

3.1.3 Other secondary endpoints

Other, secondary endpoints, according to the protocol, include the following:

- Time to first pneumonitis of grade ≥ 3 (TFP3), defined as the time from the date of enrolment until first documented pneumonitis of grade ≥ 3 . Censoring will occur at the last assessment only if patient is lost to follow-up or has died.
- Objective response, defined as best overall response (Complete or Partial Response) across all assessment time-points during the period from enrolment to termination of trial treatment.
- Time to treatment failure (TTF), defined as time from enrolment to treatment failure for any reason, including disease progression, treatment toxicity, refusal/withdrawal and death (even after treatment completion). Censoring will occur at the time of last tumour assessment only if the patient is lost to follow-up.
- Overall survival (OS), defined as the time from the date of enrolment until death from any cause. Censoring will occur at the last follow-up date.
- Toxicity, defined as adverse events classified according to CTCAE version 4.

Other exploratory endpoints that will be assessed for descriptive purposes are:

- Time to treatment discontinuation (TTD), defined as the time from enrolment to treatment discontinuation for any reason.
- Duration of response (DoR), defined as the time from documentation of tumour response (either partial or complete response) to disease progression or death.

More details on the exact definition and calculation of efficacy endpoints is provided in the supplement, Table S1, while a detailed description of AEs and SAEs is provided in a next section.

3.2 (Serious) Adverse Events

The main criterion for treatment tolerability is the occurrence of toxicities and adverse events. The severity and causality will be classified according to the NCI CTCAE Version 4. The CTCAE is available for downloading (<http://evs.nci.nih.gov/ftp1/CTCAE/About.html>).

An adverse event (AE) is defined as any untoward medical occurrence that occurs from the first dose of study medication until 30 days after the final dose, regardless of whether it is considered related to a medication. In addition, any known untoward event that occurs subsequent to the adverse event reporting period that the investigator assesses as possibly related to the protocol treatment should be considered an adverse event.

Serious Adverse Events

A Serious Adverse Event (SAE) is defined in general as any undesirable medical occurrence/adverse drug experience that occurs during or within 30 days after stopping study treatment that, at any dose, results in any of the following:

- is fatal (any cause),
- life-threatening,
- requires or prolongs inpatient hospitalization,
- results in persistent or significant disability/incapacity,
- is a congenital anomaly or birth defect,
- is a secondary malignancy,
- requires significant medical intervention.

Other significant/important medical events which may jeopardize the patient are also considered serious adverse events. Serious also includes any other event that the investigator or the ETOP Safety Office judges to be serious or which is defined as serious by the regulatory agency in the country in which the event occurred.

Severity Grade

The adverse event severity grade provides a qualitative assessment of the extent or intensity of an adverse event, as determined by the investigator or as reported by the subject. The severity grade does not reflect the clinical seriousness of the event, only the degree or extent of the affliction or occurrence (e.g. severe nausea, mild seizure), and does not reflect the relationship to study drug.

Severity grade for other adverse events not covered in the toxicity grading scale:

1=Grade 1	Mild
2=Grade 2	Moderate
3=Grade 3	Severe
4=Grade 4	Life-threatening
5=Grade 5	Fatal

Multiple recordings of an event for the same patient

There are some cases where a patient may experience the same event (AE/SAE) more than once. In such cases, the event will be counted only once for the calculation of the total number of events reported for the overall safety cohort. Nevertheless, if an SAE occurs more than once, all narratives will be included.

3.3 Analysis populations

3.3.1 Interim safety analysis cohort (applicable for previous formal safety analysis)

The cohort for the interim safety analysis includes the first 21 enrolled patients who have completed 3 months on FU post RT. Competing risks for the safety evaluation (death or treatment discontinuation before completion of a 3-month FU after end of RT) are excluded.

3.3.2 Safety cohort

The safety cohort will encompass all patients who have received at least 1 dose of trial treatment. In the present context, the main safety analysis will focus on the patients under the concomitant chemo-radiotherapy schedule, while a secondary analysis will include all patients enrolled either under the concomitant or the sequential treatment schedule.

3.3.3 Efficacy cohort

The efficacy cohort will be consisted of all eligible patients enrolled in the trial (ITT population) under AM1 and AM2 (i.e. protocol versions V2 and V3), excluding the patients that have received the sequential chemo-radiotherapy treatment schedule.

4 Study analysis

In this section, we present in detail the analyses that will be performed in the frame of final efficacy analysis.

4.1 Patient accrual & Baseline characteristics

The following information will be provided in tabular or graphical format:

- Patient accrual by center including both the enrolled and the non-enrolled patients.
- For those patients that were eventually not enrolled in the trial, an additional table will be provided, indicating the patient ID number, the center, the eligibility status (ineligible, draft, error) and the reason for non-enrolment.
- Observed vs. expected accrual will be graphically displayed.
- A consort diagram will be created to graphically depict the flow of patients and the phases of the trial.
- Summary of the baseline patient and tumour characteristics (gender, age, smoking history, ECOG performance status, histology, stage, T/N/M classification). Regarding the categorical characteristics, the frequencies of the different levels of each of them, along with the respective percentage will be presented. If missing cases exist, a separate category named “Missing” will be created. All continuous characteristics will be summarised using the following descriptive statistics: n (non-missing sample size), mean, 95% CI for the mean, median, maximum and minimum.

4.2 Treatment administration and follow-up information

Treatment information will be summarized as:

- Number of patients that started treatment
- For the chemo-radiotherapy phase:
 - Number of patients that completed chemotherapy
 - Number of patients that completed radiotherapy
 - Number of treatment failures during chemo-radiotherapy (and reasons of failure)
- For the nivolumab phase
 - Statistics (median, range) for the number of nivolumab cycles administered per patient
 - Number of complete per-protocol treatments
 - Number of nivolumab failures

- For patients who complete per protocol treatment, information on whether or not they experience a PFS event thereafter will be also provided.
- Number of patients that did not receive any dose of trial treatment, along with reasons for not doing so.
- For those patients that progressed information on further lines of treatment will be also provided.

Follow-up time of the patients will be presented as:

- Median follow-up (FU) of the patients along with the respective interquartile range (IQR) and the number (%) of patients that are still on FU. A Kaplan-Meier (K-M) will be also provided for a graphical representation of follow-up time.

4.3 Efficacy analysis

The efficacy analysis will be conducted based on the efficacy cohort and will include presentation of the following:

4.3.1 Key-secondary endpoint: One-year PFS

For the key-secondary endpoint of progression free survival (PFS) at one-year, a summary table will be produced, presenting the total number of patients evaluable at the one-year time point, the number of patients that are progression-free, along with the corresponding proportion and 90% exact binomial confidence interval. According to the statistical design, with 1-sided alpha 0.05 and power 80%, for 74 evaluable patients we need 41 patients to reach one-year progression-free (without PFS event) so as to be able to reject the null hypothesis of one-year PFS rate $\leq 45\%$ versus the alternative evaluated at PFS rate 60%. In the present frame of formal evaluation of one-year PFS, loss before a PFS event or earlier than one-year of follow-up, and never starting treatment are considered as competing risks.

4.3.2 Further efficacy analysis based on PFS

PFS will be further analysed as follows:

- The total number of PFS events observed, will be presented. In addition, 6-month, 12-month and 18-month PFS estimates, median PFS and respective 95% CIs will be provided.
- Graphical representation of PFS, will be performed via a Kaplan-Meier plot.
- To assess the effect of clinicopathological variables on PFS (gender, age, smoking history, ECOG performance status at diagnosis, histology, stage), Cox proportional hazards models will be fitted. Initially, univariate Cox models will be fitted to the data and the statistical significance of each candidate predictors will be tested at the 5% significance level (results presented in table and forest plot). Subsequently, a multivariable Cox model will be estimated, adjusted for the clinicopathological variables of interest. The backward elimination method, with a removal criterion at

10% will be used to conclude on the statistically significant variables of the model. The HRs and corresponding 95% CIs for all significant PFS predictors (in the multivariable Cox model) will be summarized in a forest plot.

- Furthermore, PFS will be analysed separately by subgroups of interest (as defined in a following section). For each subgroup number of PFS events, 6/12/18-month PFS estimates, median PFS and corresponding 95% CI as well as Kaplan-Meier plots will be presented.
- Finally, a waterfall plot will be created to graphically depict the percent tumour change, a swimmer plot to illustrate efficacy information by patient (time-on-treatment, time to response and progression, follow-up), and a spider plot to graphically overview the changes in tumour size over time.

4.3.3 Analyses of other secondary efficacy endpoints

The other time-to-event endpoints (OS, TTF, TTD, TFP3) will be estimated by the Kaplan Meier method. The median time estimations for all survival endpoints along with the respective 95% CI's, as well as the number (%) of failures (as per endpoint definition for "failure") will be provided. Additionally, for TTF, the rate of treatment failures per month of FU will be provided, derived from the total number of treatment failures over the total number of FU time (in months).

Clinical efficacy will be further described by objective response rate (ORR) as defined by RECIST v1.1. ORR will be presented along with a 95% exact binomial CI. Median DOR, along with 95% CI will be presented. Graphical representation of duration of response will be performed via Kaplan-Meier and swimmer plots.

Definition of the endpoints described in the trial protocol, as well as necessary dates, cohorts and specific clarifications are all indicated in details in the table provided in the Appendix.

4.3.4 Subgroup analysis

To determine whether the treatment effect is consistent across various subgroups the treatment effect for all efficacy endpoints will be estimated separately for each category of the following pre-specified variables.

- Histology (squamous vs non-squamous)
- Stage (IIIa vs IIIb)
- Gender

4.3.5 Exploratory analysis for patients receiving nivolumab

For patients having received at least one dose of nivolumab treatment, PFS and the other secondary efficacy endpoints (OS, TTF, TTD, TFP3, ORR, DOR) will be also evaluated having as starting point the date of first nivolumab dose.

The results that will be included in the exploratory analysis are:

- Number of PFS events, 6/12/18-month PFS estimates, median PFS and corresponding 95% CI, as well as multivariable Cox regression.
- Analogous results for OS, TTF, TTD, TFP3, DOR
- ORR with 95% exact binomial CI.

4.4 Safety analysis

The safety analysis will be based on the safety cohort and will include presentation of the:

- Overview of the number of patients who experienced an AE and/or an SAE, as well as the number of patients in the safety cohort who did not experience an event. The respective percentages will also be shown.
- Number of AEs/SAEs and rate of AE/SAE occurrence per month of FU. Again, both rates will be calculated based on the total number of AEs and SAEs, respectively, over the total number of FU time (in months).
- Number of patients experiencing a specific number of AEs/SAEs.
- Distribution of (S)AEs by grade and CTCAE category. Six columns, one for each grade and one for all (any) grades, will be shown. An additional column indicating which events were SAEs -or started as AEs and became SAEs later on- will be available. In this column, the frequency of the SAEs and the severity grade will be given. The percentages that will accompany the frequencies will be based on the respective frequency of an event over the total number of patients in the safety cohort. This table will include all (S)AEs irrespective of their relation to the trial treatment.
- Analogous table focusing only on the treatment related (S)AEs.
- Number and corresponding percentages of treatment related (S)AEs by grade, leading either to treatment discontinuation or death
- Distribution of the maximum severity grade experienced by the patients
- For all fatal SAEs, cause of death will be provided.
- For all pneumonitis of grade ≥ 3 , detailed, by patient, information for the event (start/end date, outcome) as well as for the radiotherapy administered (event start date compared to radiotherapy end date, total dose, mean lung dose, V20 lung dose) will be provided. For comparison purposes, descriptive statistics of the mean lung dose (mean with 95% CI and median with IQR) will be provided for patients with pneumonitis of lower grade (1-2) or without such an event.
- Analogous by-patient tables for all dyspnoea of grade ≥ 3 , including information on mean heart dose, as well as for all esophagitis of grade ≥ 3 , including oesophageal mean dose.

- Results of logistic regression models, modelling the occurrence of adverse events of primary interest (pneumonitis, dyspnoea, esophagitis) by grade and adjusting for the mean dose of radiation.

4.5 Missing Data

- Baseline characteristics

For categorical baseline characteristics if missing cases exist, a separate category named “Missing” will be created. As far as continuous values, missing cases will not be replaced by any statistics calculated over non-missing data.

- Dates:

If the day of the month is missing for any date used in the analysis, the 15th of the month will be used to replace the missing date unless the calculation results in a negative time duration (e.g., date of onset cannot be prior to day one date). In this case, the date resulting in one day of duration will be used. If the day of the month and the month are missing for any date used in a calculation, i.e., January 1 will be used to replace the missing date. Missing dates for adverse events will be imputed based on the similar principle.

- Incomplete tumour assessment information

In patients who have no on-study assessments:

- If death is recorded prior to the first planned tumour assessment, the death date will be considered as the date of the PFS event.
- If clinical progression is recorded prior to the first planned tumour assessment, the date of the reported clinical progression will be considered as the date of the PFS event.

4.6 Sensitivity analysis

In a sensitivity analysis framework, the efficacy analysis will be repeated using the safety, instead of efficacy, cohort.

In particular, the results that will be included in the sensitivity analysis are:

- Number of PFS events, 6/12/18-month PFS estimates, median PFS and corresponding 95% CI, as well as multivariable Cox regression.
- Analogous results for OS, TTF, TTD, TFP3, DOR
- ORR with 95% exact binomial CI.

4.7 Presentation of results

The results will be presented through tables and figures. A summary of the results will also accompany the main report. First, a short synopsis of the results will be presented through bullets, where only the most important findings will be shown. Following that, a more detailed description of the results will be provided, sectioned in the following order:

- Patient accrual and baseline characteristics
- Treatment administration
- Efficacy analysis
- Safety analysis

All tables and figures will be included in an appendix.

5 Technical details

Data will be primarily analysed using the SAS software package (version 9.4 or later), while the R statistical software will be also used for specific analyses and plots.

A second statistician, the reviewing statistician, will independently reproduce all analysis and summary statistics. The reviewing statistician will have an overview of the entire analysis and will explicitly check the code producing tables and figures, as well as any other pieces of code as desired.

Reporting conventions

P-values ≥ 0.001 will be reported to 3 decimal places; p-values > 0.010 will be reported with two significant digits; p-values less than 0.001 will be reported as “ < 0.001 ”. The mean, 95% confidence limits, quantiles, and any other statistics, will be reported to one decimal. Hazard ratios (HRs) and their 95% CI's will be reported to two decimals. Estimated parameters, not on the same scale as raw observations (e.g. regression coefficients) will be reported to 3 significant figures.